The Maximum Edge Biclique Problem is NP-Complete

Peeters, M.J.P.

Publication date:
2000

The maximum edge biclique problem is NP-complete

René Peeters
Tilburg University

May 9, 2000

Abstract

We prove that the maximum edge biclique problem in bipartite graphs is NP-complete.

A biclique in a bipartite graph is a vertex induced subgraph which is complete. The problem of finding a biclique with a maximum number of vertices is known to be solvable in polynomial time but the complexity of finding a biclique with a maximum number of edges was still undecided.

1 Introduction

Let $G = (V, E)$ be a graph with vertex set V and edge set E. A pair of two disjoint subsets A and B of V is called a *biclique* if $\{a, b\} \in E$ for all $a \in A$ and $b \in B$. Thus the edges $\{a, b\}$ form a complete bipartite subgraph of G (which is not necessarily an induced subgraph if G is not bipartite). A biclique $\{A, B\}$ clearly has $|A| + |B|$ vertices and $|A| * |B|$ edges. In this note we restrict ourselves to case that G is bipartite. The two colour classes of G will be denoted by V_1 and V_2, so $V = V_1 \cup V_2$.

Already in the book of Garey and Johnson [2] (problem GT24) the complexity of deciding whether or not a bipartite graph contains a biclique of a certain size is discussed. If the requirement is that $|A| = |B| = k$ for some integer k (this is called the *balanced complete bipartite subgraph problem* or *balanced biclique problem*), then the problem is NP-complete. If however the requirement is that $|A| + |B| \geq k$ (the *maximum vertex biclique problem*), the problem can be solved in polynomial time via the matching algorithm. The complexity of the maximum vertex biclique problem for general graphs depends on the precise definition of a biclique in this case. With the above definition the problem is solvable in polynomial time since there is a one to one correspondence between bicliques in the bipartite double of the graph and bicliques in the graph itself. If one defines a biclique as an induced complete bipartite subgraph (so A and B are independent sets in G), then the maximum vertex biclique problem for general graphs is NP-complete (see [10]). A natural third variant is the so-called *maximum edge biclique problem (MBP)* where the requirement is that $|A| * |B| \geq k$. Up to now the complexity of this problem was still undecided.

In various papers the complexity of MBP is mentioned and guessed to be NP-complete ([1, 4, 3, 9]. In [1] some applications of MBP are discussed and it is shown that the weighted version of MBP is NP-complete. Furthermore the authors show that four variants of MBP are NP-complete. Using different techniques Hochbaum [4], Haemers [3] and Pasechnik [9] derive upper bounds for the maximum number of edges in a biclique of a bipartite graph. Hochbaum [4] presents a 2-approximation algorithm for the minimum number of edges needed to be removed so that the remainder is a biclique based on an LP-relaxation. Inspired by the work of Lovász on the Shannon capacity of a graph ([6]), Haemers [3] and Pasechnik [9] derive similar inequalities for the maximum biclique problem. Pasechnik uses semidefinite programming techniques wereas Haemers uses eigenvalue techniques.

In the next section we prove that indeed MBP is NP-complete. The reduction used is inspired by the reduction that is used to prove the NP-completeness of the balanced biclique problem (see [5]). As a consequence MBP is also NP-complete for general graphs.

2 The reduction

We define MBP as follows:

Maximum edge biclique problem (MBP): Given a bipartite graph $G = (V_1 \cup V_2, E)$ and a positive integer K, does G contain a biclique with at least K edges?

Theorem 1 *MBP is NP-complete.*

Proof: We reduce 3SAT to MBP in two steps. Given an instance ϕ of 3SAT, we first construct a graph $G = (V, E)$ that has a clique of size $\frac{1}{2}|V|$ if and only if ϕ is satisfiable. This reduction is a modification of a well known and rather straightforward reduction from 3SAT to CLIQUE/INDEPENDENT SET ([7, 8]). Secondly we construct a bipartite graph $H = (V_1 \cup V_2, E')$ such that H has a biclique containing a certain number of edges if and only if G has a clique of size $\frac{1}{2}|V|$. This second step is a modification of the reduction from CLIQUE to BALANCED COMPLETE BIPARTITE SUBGRAPH referred to in [2] (problem GT24) and published in [5].

We are given an instance ϕ of 3SAT with m clauses $C_1, \ldots, C_m$, with each clause being $C_i = (\alpha_{i1} \vee \alpha_{i2} \vee \alpha_{i3})$, with the α_{ij}'s being either Boolean variables or negations thereof. Now construct the graph $G = (V, E)$ as follows:

$$\begin{aligned}
V &= \{v_{ij} : i = 1, \ldots, m; j = 1, 2, 3\} \cup \{v_i : i = 1, \ldots, m\} \\
E &= \{\{v_{ij}, v_{kl}\} : i \neq k; \alpha_{ij} \neq \neg\alpha_{kl}\} \\
&\cup \{\{v_{ij}, v_k\} : i = 1, \ldots, m; j = 1, 2, 3; k = 1, \ldots, m\} \cup \{\{v_i, v_j\} : i \neq j\}
\end{aligned}$$

Clearly a maximal clique in G contains all vertices v_i and at most one vertex out of each triple $\{v_{i1}, v_{i2}, v_{i3}\}$. It is easy to check that G has a (maximal) clique of size $2m$ ($=\frac{1}{2}|V|$) if and only if ϕ is satisfiable.

Let $k = \frac{1}{2}|V|$. Now construct an instance $H = (V_1 \cup V_2, E')$, K of MBP as follows: Let

$$\begin{aligned}
V_1 &= V \\
V_2 &= E \cup \{e_1, \ldots, e_{\frac{1}{2}k^2 - k}\} \\
E' &= \{\{v, e\} : v \in V; e \in E; v \notin e\} \cup \{\{v, e_i\} : v \in V; i = 1, \ldots, \frac{1}{2}k^2 - k\} \\
K &= k^3 - \frac{3}{2}k^2
\end{aligned}$$

This construction can clearly be performed in polynomial time. Suppose G has a clique C of size k. Take $A := V - C$ and $B := \{e_1, \ldots, e_{\frac{1}{2}k^2 - k}\} \cup \{\{c, d\} : c, d \in C; c \neq d\}$. Then $\{A, B\}$ is a biclique with $|A| * |B| =$

$k*(\frac{1}{2}k^2 - k + \frac{1}{2}k(k-1)) = k^3 - \frac{3}{2}k^2$. So if G has a clique of size k then H has a biclique with $k^3 - \frac{3}{2}k^2$ edges. On the other hand, if H has a biclique with at least $k^3 - \frac{3}{2}k^2$ edges, then G must have a clique of size k. We complete the proof by showing this.

Let $\{A, B\}$ be a biclique of H with $A \subseteq V_1$ and $B \subseteq V_2$. We shall prove that $|A| * |B| \leq k^3 - \frac{3}{2}k^2$ and that equality implies that G has a clique of size k. Without loss of generality, $e_i \in B$ for $i = 1, \ldots, \frac{1}{2}k^2 - k$. Let $a := |A|$ and $b := |B| - (\frac{1}{2}k^2 - k)$.

The b elements of $B \cap E$ correspond with edges in G whose endpoints are not in A. There are $2k - a$ vertices of G that are not in A so $b \leq \frac{1}{2}(2k-a)(2k-a-1)$, with equality if and only if $V - A$ is a clique with edge set $B \cap E$.

We consider two cases:

1. Suppose $a \geq k$, so $|V - A| = k - c$ with $c := a - k$ (So $0 \leq c \leq k$). Then $b \leq \frac{1}{2}(k-c)(k-c-1)$, so

$$|A| * |B| \leq [k + c] * [\frac{1}{2}k^2 - k + \frac{1}{2}(k-c)(k-c-1)]$$

   This reduces to

$$|A| * |B| - (k^3 - \frac{3}{2}k^2) \leq \frac{1}{2}c(c^2 - (k-1)c - 2k)$$

   Now $c^2 - (k-1)c - 2k$ is negative for $0 \leq c \leq k$, so $|A| * |B| \leq k^3 - \frac{3}{2}k^2$ with equality if and only if $|A| = k$ and $V - A$ is a clique of size k in G.

2. Suppose $a \leq k$, so $|V - A| = k + c$ with $c := k - a$ (So $0 \leq c \leq k$). Since G has no cliques with more than k vertices, the number of edges b in the subgraph of G induced by $V - A$ is at most $\frac{1}{2}(k+c)(k+c-1) - c$. This leads to

$$|A| * |B| \leq [k - c] * [\frac{1}{2}k^2 - k + \frac{1}{2}(k+c)(k+c-1) - c]$$

   This reduces to

$$|A| * |B| - (k^3 - \frac{3}{2}k^2) \leq \frac{1}{2}c^2(-c + 3 - k)$$

Since we may assume that $k \geq 4$, again the right hand side is negative for $1 \leq c \leq k$ and zero for $c = 0$. So $|A| * |B| \leq k^3 - \frac{3}{2}k^2$, with equality if and only if $|A| = k$ and $V - A$ is a clique of size k in G.

$\square$

References

[1] M. Dawande, P. Keskinocak and S. Tayur, *On the biclique problem in bipartite graphs*, GSIA Working Paper 1996-04, Carnegie Mellon University, Pittsburgh, PA 15213, USA, 1996.

[2] M.R. Garey and D.S. Johnson, *Computers and Intractability, A guide to the theory of NP-completeness*, Freeman, San Francisco, 1979.

[3] W.H. Haemers, Bicliques and eigenvalues, Research Memorandum FEW 783, Tilburg University, 1999.

[4] D.S. Hochbaum, Approximating clique and biclique problems, *J. Algorithms*, 29(1): 174-200, 1998.

[5] D.S. Johnson, The NP-completeness column: an ongoing guide, *J. Algorithms*, 8(3): 438-448, 1987.

[6] L. Lovász, On the Shannon capacity of a graph, *IEEE Trans. Inform. Theory*, 25: 1-7, 1979.

[7] K. Mehlhorn, *Data Structures and Algorithms 2: Graph Algorithms and NP-Completeness*, Springer-Verlag, Berlin, 1984.

[8] C.H. Papadimitriou, *Computational Complexity*, Addison-Wesley, 1994.

[9] D.V. Pasechnik, *Bipartite sandwiches: bounding the size of a maximum biclique*, preprint, 1999.

[10] M. Yannakakis, Node- and edge-deletion NP-complete problems, *Proc. 10th Ann. ACM Symp. on Theory of Computing*, Association for Computing Machinery, New York, 253-264, (1978).